Chapter 5

AIR POLLUTION AND EARNINGS: EMPIRICAL RESULTS

Standard least-squares estimating techniques are readily applied to the second stage (the stage in which the individual has decided to do at least some picking during a day) of the representation in the last chapter of the picker's decision problem. Empirical implementation of this representation requires information on earnings, the piece-work wage rate, boxes of fruit picked, picker and crew hours worked, and grove and environmental attributes. The objective is to estimate, with respect to changes in air pollution, the changes in the individual picker's earnings when he is picking fruit and the changes in the leisure he voluntarily takes when he could have been picking fruit. The maintained hypothesis is that picker earnings vary negatively, and voluntarily taken hours of leisure vary positively, with respect to increases in air pollution. This chapter deals with the impact of air pollution upon earnings. The next chapter presents results for absenteeism.

Throughout the empirical investigation described in this and the succeeding chapter, the overriding criterion has been to arrive at estimated expressions yielding unbiased, or at least consistent, coefficients for a particular explanatory variable, the air pollution variable. Consistent estimates of these coefficients allow us to make inferences about the compensation in terms of earnings the picker requires to make him indifferent to an increase in the level of air pollution. They can also enable us to infer the change in the picker's equilibrium hours worked with respect to changes in air pollution.

One common source of bias in estimated coefficients is "data-grubbing." In the words of Selvin and Stuart (1966, p. 21):

> " . . . any preliminary search of data for a model, even when the alternatives are predesignated, affects the probability levels of all subsequent tests based on that model on the same data, and in no very simple way, and also affects the characteristics of subsequent estimation procedures. The only valid course is to use different data for testing the model dredged from the first set of data."

In order to assure that the data used to test the hypotheses generated from the analytical framework of the previous chapter are unsullied by any prior efforts to put together an empirical model from a combination of analytical and empirical investigations, the picking histories of four experienced pickers who worked more-or-less continuously picking lemons over an entire year were used for preliminary estimation. All "data-grubbing" for the empirical results presented in this report was limited entirely to these four pickers; that is, the estimates for all pickers other than the four issue from virgin data.

Estimates of the Inverse Supply Function. The analytical framework surrounding Figure 4.2 makes it convenient to estimate the inverse supply function, the function in which earnings are determined by hours worked. The coefficient attached to the air pollution variable in an empirical counterpart:

I_t = f (H, air pollution, other factors that shift S in Figure 4.2(b))

to the S function of Figure 4.2(b) will, for the sample observations on the picker's work performance, then measure the shift of S averaged over all the hours the individual picker worked.

For estimation purposes, several assumptions, in addition to those embodied in the analytical framework, were imposed upon the stochastic form of the above earnings expression. First, in the absence of unique directions about functional form from the picker's decision model, all estimated expressions have been specified in double-logarithmic form. Although no formal comparisons were made, exploratory manipulations with the four preliminary test pickers made it appear that the double-log form fit each picker's data equally as well as arithmetic or semi-log forms. The double-log form was ultimately selected because of its greater flexibility. In particular, it permits the marginal effect of air pollution upon individual -picker earnings to be constant, decreasing, or increasing; it makes the coefficients of the explanatory variables easily interpretable as constant elasticities; it restricts the dependent variables to positive values; it reduces the influence of extreme data values; and, finally, it may reduce heteroskedasticity,

Second, the picker was always assumed to work the same number of hours as his crew. Thus only those observations in which the picker's work-day was the same length as that of his crew were used for estimation purposes. This implies that the picker does not view his hours-worked as a decision variable, but rather accepts H as exogenously determined by the crew foreman through the, foreman's choice of H^+. This is equivalent to assuming that Z = 0 in (12). The assumption was adopted for two reasons: (1) to providc an estimate of the response of earnings to air pollution independent of states in which nonzero combinations of earnings and leisure could be chosen; and most importantly, (2) to reduce possibilities of introducing bias into the estimated air pollution coefficients because of failure to include a variable relevant to explanation of the variations in the picker's earnings. As will be shown in a later chapter dealing with absenteeism, our data does not permit us to account for very much of the variation in the hours the picker chooses to work. Since hours-worked is an integral element of the earnings expression, an inability to explain very much of the variation in hours-worked substantially increases the risk that a relevant variable, nonorthogonal to the air pollution variable, is being neglected. The assumption,that hours-worked is exogenous rather than endogenous to the picker's problem neatly avoids this.

Third, the variable representing hours-worked could, because of the common practice in the packinghouse crew records of rounding to the nearest half hour, contain a relatively high degree of measurement error when the picker worked for only a short time in a given grove. In order to correct for this possible source of error, we assumed for the earnings expression that the picker always worked at least two hours when he worked at all. The adoption of this assumption required that all observations in which the picker worked less than two hours be excised from the data used for estimation purposes.

Finally, an examination of the residual pattern of some ordinary-least-squares regressions for the four preliminary test lemon pickers revealed a definite drift of the residuals across time, even though each of these four individuals were known to have been picking lemons for years. An obvious means of ameliorating this is to introduce a calendar date variable into the regression specification. This additional variable might capture the work performance effects of selective picking as opposed to clean picking of lemon groves. Workers engaged in lemon picking are required to use rings and pick by color during most of the multiple harvests in a lemon grove during a calendar ycor. Inclusion of a variable representing calendar date mny capture the effect of prior picking that has occurred in a grove or it'may register factors not explicitly recorded that do influence picking ease, Since all orange picking is clean picking, calendar date did not seem relevant. Inspection of the residuals for ordinary-least-squares estimates of an orange picker's earnings expression seemed to confirm this irrelevancy.

Table 5.1 below presents the acronyms of the variables used to estimate the earnings expression. The initial ordinary-least-squares estimates of the earnings expression for the four preliminary test pickers revealed several additional statistical problems. As has been stressed in previous pages, our primary concern was to obtain consistent estimates of the coefficient for the air pollution variable. However, if it is not statistically confirmed that the air pollution variable is significantly different from zero at the usual test levels, then any claims as to the effect of air pollution upon the performance, of citrus pickers is unfounded. As is well known, collinearity inflates the standard errors of the set of explanatory variables. In turn, this implies a reduction in the t-statistics and an unnecessarily conservative test of significance. Different equation specifications for the four preliminary test pickers verified that TM and DE were relevant explanatory variables and consequently must be included in the empirical specification. Additionally, both were collinear with the air pollution variables and thus rendered difficult an interpretation of the levels of statistical significance of these variables.

Tables 5.2 and 5.3 provide detail on the extent of collinearity between the air pollution variables and temperature. Each picker is identified by the general locale in which he picked fruit as well as by'a number immediately following the locale. The parenthetic numbers indicate the years to which data for the picker refer. Thus Upland 1 (1973) refers to the same individual as Upland 1 (1974), but the year from which the data is drawn is different. Those to whom we refer to as pickers are thus on occasion the same individual distinguished by year and/or crop. A distinction was made between crop years for the picking activities of the same individual because the time pattern of the fruit harvest is said by growers and packinghouses to have differed fairly substantially between 1973 and 1974.[1]

Table 5.1

Glossary of Variable Names

B - The number of 3115 cubic inch field boxes picked by the picker during the work-day in a particular grove.

BT - The average number of 3115 cubic inch field boxes picked per tree during the work-day by the picker's crew in a specific grove.

DE - The calendar date of the work day: Jan. 1 = 1; Dec. 31 = 365.

I_t - The picker's daily gross earnings in dollars from picking activities for each grove worked.

I_{t-1} - The picker's gross earnings in dollars from picking activities in the previous pay period.

H - The number. of hours the picker spent in picking activities in each grove.

FR - The number of fruit from the grove required to fill a 3115 cubic inch field box.

ØZ - The arithmetic average 24-hour ambient concentration on the work day of 0 in parts per million by volume as measured by the CHEMILUM method a2 the monitoring station closest to the grove site.

ØZH - The arithmetic average of the hourly ambient concentrations of O_3 occurring during the time interval the worker was actually engaged in citrus picking. This variable is also measured by the CHEMILLJM method at the monitoring station closest to the grove site.

TM - The maximum hourly arithmetic average dry-bulb temperature in F^o on the work-day at the monitoring station closest to the grove site.

TR - An index indicating the height of the trees picked by the worker's crew during the work day.
1 = tree can be picked without a ladder.
2 = ladder picked trees up to 9 1/2 feet tall.
3 = ladder picked trees 9 1/2 to 12 feet tall.
4 = ladder picked trees in excess of 12 feet tall.

W - The rate-of-pay (in dollars x 10) the picker receives for each 3115 cubic inch field box of citrus he picks.

Table 5.2

Simple Correlation Coefficients Between ØZ, ØZH, and TM for Various Lemon Pickers.

Worker r:	ØZ.ØZH	ØZ.TM	ØZH.TM
Upland 1 (1973)	.735	.744	.588
Upland 1 (1974)	.643	.753	.555
Upland 2 (1973)	.737	.749	,590
Upland 2 (1974)	.658	.747	.557
Upland 3 (1973)	,711	.749	,593
Upland 4 (1973)	.752	.835	.653
Upland 22 (1974)	.558	.753	.554
Santa Paula 10 (1973)	N.A.	.478	N.A.
Santa Paula 10 (1974)	N.A.	.486	N.A.
Santa Paula 11 (1973)	N.A.	,472	N.A.
Santa Paula 11 (1974)	N.A.	.453	N.A.

N.A. - Not Available.

Table 5.3

Simple Correlation Coefficients Between ØZ, ØZH, and TM for Various Orange Pickers

Worker r:	ØZ.ØZH	ØZ.TM	ØZH.TM
Upland 2 (1973)	N.A.	.688	N.A.
Upland 4 (1973)	N.A.	,703	N.A.
San Bernardino 5 (1973)	.824	.763	.654
San Bernardino 7 (1973)	.747	.751	.585
Irvine 38 (1974)	N.A.	.223	N.A.
Irvine 39 (1974)	N.A.	.205	N.A.
Irvine 40 (1974)	N.A.	,223	N.A.

N.A. - Not Available

For both lemon and orange pickers, the collinearity between air pollution and temperature is somewhat less for ØZH than for ØZ. Whenever available data allow it, the former measure of air pollution is used in the estimated expressions. Some collinearity between the temperature and air pollution variables was, of course, expected. The standard-way to resolve this complication is to obtain an unbiased estimate of the temperature coefficient from an extended sample, or from a sample where the correlation between the collinear variables is less severe, and then to subtract the temperature term from the dependent variable, thus forming a new regression specification. Because air pollution and temperature in the Ventura region do not appear to be highly correlated, it would seem that data collected from that region would be ideal for this purpose. However, after considerable consternation, we decided not to follow this procedure. Our reasons were two. First, it was felt that the transformation of the dependent variable would alter the interpretation of the estimated coefficients. It is not clear to us, in terms of the analytical framework of Chapter 4, what the use is of a measure , of the compensating surplus for air pollution where the picker has already received the compensation required to make him indifferent between existing temperatures and some hard-to-identify temperature he regards as ideal for his citrus fruit picking activities. Second, and perhaps more important, the temperature coefficient used to transform the earnings variable for one individual would, of necessity, be the coefficient estimated for another individual or set of individuals. It is generally acknowledged that responses to identical perturbations in meterological and environmental variables can differ greatly across individuals. The possibility of introducing bias into the other estimated coefficients, particularly the air pollution coefficients, therefore seemed, in our judgment, to be excessive, We thus instead chose to present regression results where the temperature variable is both included and excluded, leaving it to the reader to judge for himself where the "true" level of significance lies.

With respect to the calendar date variable, we attempted to mollify the collinearity problem by viewing the system as recursive. Specifically, we hypothesized the following pair of expressions:

(15a) $L\hat{w} = f(LBT, LFR, LTR, LDE)$,

(15b) $LI_t = g[L\hat{w}, LH, L\text{ (air pollution)}, LTM, L\text{ (grove attributes)}]$

where L denotes the natural logarithm of the original arithmetic value of the variable. (15a) is interpreted as the piece-work wage rate faced at the beginning of each day by the crew of the individual picker, and, given that the individual has chosen to pick during that day, (15b) is the picker's actual earnings expression. The expression of empirical interest, (15b), can be estimated by the two-stage-least-squares method. Assuming the usual classical conditions for the general linear model hold, consistent estimates of coefficients for the explanatory variables will be obtained.

The adoption for lemon pickers of the system represented by (15a) and (15b) introduced an additional collinearity problem. In particular, a linear combination of the grove attributes included as arguments in the earnings expression is highly collinear with the estimated rate-of-pay variable also appearing on the right-hand side of (15b). Theoretically, the inclusion of all these variables is required; but a specification of this sort reduces the rank of the data matrix below that required for satisfaction of the order condition for identification. Consequently, it was necessary to delete one of the grove attribute variables from (15b). The correlations of the various grove attribute variables with the air pollution variables served as our principal guide in determining the best variable to delete from the second structural equation. It cnn be shown that exclusion of a potentially relevant but orthogonal variable will not bias the air pollution coefficient, although it will increase the standard error. For the four preliminary test pickers, a review of the simple correlation coefficients revealed that the tree height variable, TR, was relatively uncorrelated (approximately -.06) with the air pollution variable for three of the pickers. For one picker the simple correlation between the two variables was high; however, it was also positive, even though the citrus industry universally expects, for given piece-work wage rates, that tree height and earnings per grove picked will vary inversely, These two facts for this single picker (the high simple correlation and its positive sign), imply that the bias imparted to the air pollution coefficient of this picker would be negative, All these considerations for the four preliminary test pickers led us to delete TR for all pickers in all two-stage-least-squares regression specifications.

Apart from the variables explicitly introduced through the theoretical model, conversations with various labor'camp managers produced suggestions about potentially relevant factors. It appears to be a widely held notion that the performance of the typical citrus harvest worker will markedly decrease on Fridays and Mondays. Expectations of a wild weekend and supposed fulfillment of these expectations are offered as a rationale for this occurrence. For the four workers initially surveyed, however, estimated coefficients for dummy (0 ,1) variables for Friday and Monday did not yield estimates significantly different from zero at the usual levels. Consequently, these variables were not included in our final specification.[2]

A second common observation is that many pickers set an earnings goal and will not work as productively once this goal is achieved. Since pickers receive weekly paychecks, one way of ascertaining the validity of this hypothesis is to include a measure of the worker's total earnings in previous weeks. Inclusion of such a variable in the two-stage-least-squares formulation for the four preliminary test workers did not yield statistically significant estimates. Hence, this variable was also not included in the estimated expressions for other workers.

Finally, labor camp managers believe that multiple *groves* worked in a day seriously impairs the productivity of the worker. It is thought that moving three or four times a day causes the worker to go through three or four "warm up" periods, thus slowing down his picking output. Again, inclusion of a single explanatory variable representing number of groves picked did not yield a coefficient significantly different from zero.

It should be mentioned that -although none of the three variables alluded to above proved to be statistically significant, the expected signs were in fact obtained,

Wc have referrcd several times to "thc air pollution variablcs" in the preceding discussion without explicitly stating in each circumstance what measure we arc postulating. Idcolly , one would like thc pollution monitoring stations to be located in each and every grove, with hourly readings having a One-to-one correspondence with hourly picking performance. Unfortunately, the picking data cannot be disaggregated to this extent. Moreover, we also must rely on the hourly readings from the closest recording station.

In some instances, we were unable to obtain the start and stop times for the picking crews for the various groves, and in these cases we employed the arithmetic average 24-hour ambient concentration of air pollution in parts per million by volume as measured by the CHEMILLJM method. For the majority of workers for whom we were able to ascertain the actual time period of the day during which they picked fruit, we used the arithmetic mean of the hourly concentrations during this time span as our measure of air pollution.

Several alternative characterizations of the air pollution measure can be hypothesized. These might include higher moments (e.g., the variance) or perhaps some distributed lag structure. We have not attempted a distributed lag specification, but inclusion of the variance of the air pollution measure has not proven to be significant in various trial runs for the four preliminary test pickers.

The reader is by now no doubt aware that we are aware of the likely existence of measurement error in the air pollution variable. Since it does not appear possible to identify any systematic deviations in the values of this variable, it would seem that an instrumental variable would be our best recourse. The most likely candidate for an instrument, among those variables available to us, would be temperature, TM. However, this variable is already included in the regression specification. The next best alternative was posited to be TM lagged one period, A sample run for two of our trial pickers using the maximum temperature of the previous day as a proxy for the actual air pollution during the period of the next day in which the worker picked did not produce any interesting results. Given that the correlation between this lagged temperature and air pollution was only about 0.70, little gain from this reformulation could be expected. In the results to be presented below', the air pollution variable itself is utilized.

This completes the description of the basic model specifications used for estimation. The empirical specifications finally settled upon for the four preliminary test pickers were carefully checked for conformity with the classical linear model. An effort was, in fact, made to employ a program package for the formal Ramsey (1969) tests for specification error. Since time and available resources did not permit the correction of package programming errors, resort was had to less formal means. Heteroscedastic disturbances were searched for by evaluating scatter diagrams of residuals versus values of the dependent variable. No heteroscedasticity appears to be present in the final estimated expressions for the four preliminary test pickers. All variables for which data was available and which might plausibly be nonorthogonal to the air pollution variables were included at one time or another in specifications for the preliminary test workers. All those not found wanting in terms of statistical significance were included in all subsequent specifications. Autocorrelation was evaluated by means of the Durbin-Watson statistic. In the multiplicative form of the earnings expression finally selected, checks were made to assure that the disturbances were at least approximately log-normally distributed.

This completes the description of how we arrived at the basic specifications used for estimation of the earnings expression. Estimates for these expressions appear in Tables 5.4 and 5.5. Note that with one statistically insignificant exception for lemon pickers and two statistically insignificant exceptions for orange pickers, the coefficients for the air pollution variables are consistently of the expected sign. Moreover, even though a collinear temperature variable is included, the coefficients for almost half (seven of eighteen) the pickers are statistically significant at traditional levels for non-rejection of the maintained hypothesis that air pollution has a detrimental influence upon picker earnings. An indication of the impact that collinearity between the air pollution and temperature variables has upon the statistical significance of the former can be obtained from Table 5.5, where the temperature variable has been deleted from the empirical specification. In Table 5.5, all but two air pollution coefficients (both for orange pickers) have the expected negative sign and these negative coefficients are statistically significant for thirteen of the eighteen pickers. A further comparison of Table 5.5 with Table 5.4 makes it appear that the impact of the deletion of the temperature variable upon the statistical significance of the air pollution coefficients in Table 5.5 was greater for those pickers having nonsignificant air pollution

Table 5.4A

Earnings Estimates by Two-Stage-Least-Squares for Lemon Pickers. Dependent Variable = LI_t.

Picker: Variable	Upland 1[a] (1973)	Upland 1 (1974)	Upland 2[a] (1973)	Upland 2 (1974)	Upland 3 (1973)	Upland 4[b] (1973)	Upland 22 (1974)	Santa Paula 10[a] (1973)	Santa Paula 10 (1974)	Santa Paula 11[a] (1973)	Santa Paula 11 (1974)
Constant	1.326	1.270	3.180	4.713	0.831	-2.220	0.214	0.931	0.842	3.800	1.650
	(1.680)	(1.210)	(1.740)	(1.306)	(1.502)	(1.570)	(1.700)	(2.588)	(0.523)	(1.490)	(0.495)
Lw	-0.207	-0.780	-0.475	-0.509	-0.604	1.610	0.830	-0.270	-0.122	-0.473	-0.320
	(0.472)	(0.423)	(0.479)	(0.404)	(0.493)	(0.409)	(0.594)	(0.604)	(0.163)	(0.889)	(0.122)
LH	1.084	1.080	1.010	1.028	1.021	0.960	1.074	0.986	1.070	1.150	1.060
	(0.041)	(0.039)	(0.046)	(0.043)	(0.043)	(0.047)	(0.049)	(0.103)	(0.023)	(0.055)	(0.021)
LFR	0.073	0.268	-0.092	-0.129	0.255	0.107	-0.199	0.388	0.068	-0.002	0.120
	(0.297)	(0.219)	(0.342)	(0.221)	(0.359)	(0.329)	(0.349)	(0.354)	(0.129)	(0.289)	(0.103)
LBT	-0.018	-0.204	0.113	-0.067	-0.070	0.339	0.243	-0.314	-0.119	-0.155	-0.106
	(0.108)	(0.114)	0.108	(0.111)	(0.359)	(0.098)	(0.162)	(0.102)	(0.040)	(0.145)	(0.026)
LOZ								-0.242**	-0.034***	-0.029	0.002
								(0.128)	(0.016)	0.074	(0.019)
LOZH	-0.015	-0.023	-0.038**	-0.017	-0.029*	-0.012	-0.047*				
	(0.022)	(0.024)	(0.021)	(0.026)	(0.018)	(0.018)	(0.031)				
LTM	-0.173	0.028	-0.164	-0.414***	-0.085	-0.178	-0.037	-0.266	-0.042	-0.351	-0.142***
	(0.175)	(0.148)	(0.186)	(0.158)	(0.178)	(0.172)	(0.190)	(0.455)	(0.073)	(0.295)	(0.070)
$\bar{R}^2$	0.824	0.856	0.824	0.822	0.849	0.888	0.822	0.640	0.877	0.927	0.912
S.E.	0.246	0.209	0.287	0.206	0.217	0.176	0.257	0.304	0.169	0.174	0.155
D-W	1.88	2.09	1.79	1.86	1.92	1.70	1.82	1.68	1.94	1.72	1.57
F	131.46	153.18	143.24	120.52	127.87	124.86	104.74	15.11	381.42	99.92	456.11
Sample Size	168	162	189	156	136	101	143	58	293	54	264
Sample Period	March 17- Dec. 21	April 1- Nov. 2	March 19- Dec. 21	April 1- Nov. 2	March 17- Dec. 20	March 16- Dec. 21	April 17- Nov. 2	May 14- July 11	Jan. 3- Nov. 20	May 13- July 11	March 3- Dec. 4

Table 5.4A
(continued)

[a]The results reported for this picker have been derived from various empirical testing procedures and regression specifications.

[b]This picker is a woman.

Levels of significance are explicitly shown for only the air pollution and temperature variables.

*Coefficient is significantly different from zero at the 0.10 level of the one-tailed t-test.

**Coefficient is significantly different from zero at the 0.05 level of the one-tailed t-test.

***Coefficient is significantly different from zero at the 0.025 level of the one-tailed t-test.

(The numbers in parentheses are the standard errors of the estimated coefficients).

Table 5.4B

Earnings Estimates by Ordinary-Least-Squares for Orange Pickers.[b] Dependent Variable = LI_t.

Picker: Variable	Upland 2 (1973)	Upland 4 (1973)	San Bern. 5 (1973)	San Bern. 7 (1973)	Irvine 38 (1974)	Irvine 39 (1974)	Irvine 40 (1974)
Constant	2.210	2.094	0.898	-0.786	3.583	1.279	2.387
	(2.062)	(1.325)	(1.285)	(1.363)	(1.190)	(1.282)	(1.175)
Lw	0.290	0.323	0.119	1.189	a	a	a
	(0.288)	(0.181)	(0.178)	(0.252)			
LH	1.371	1.114	0.884	1.029	1.001	1.146	1.172
	(0.109)	(0.686)	(0.077)	(0.090)	(0.067)	(0.062)	(0.066)
LTR			-0.162	0.082			
			(0.112)	(0.167)			
LBT			-0.043	0.150			
			(0.051)	(0.095)			
LØZ	0.065	-0.009			-0.081*	-0.027	0.065
	(0.075)	(0.047)			(0.059)	(0.062)	(0.057)
LØZH			-0.061***	-0.054			
			(0.030)	(0.053)			
LTM	-0.567	-0.468**	0.028	-0.226	-0.568"	-0.096	-0.366*
	(0.043)	(0.276)	(0.283)	(0.312)	(0.270)	(0.293)	(0.265)
$\bar{R}^2$	0.755	0.864	0.487	0.814	0.692	0.759	0.756
S.E.	0.181	0.110	0.410	0.397	0.308	0.332	0.303
D-W	1.276	2.043	1.800	1.502	1.328	1.557	1.111
F	44.094	77.782	24.724	127.562	82.392	116.543	113.519
Sample Size	57	54	163	152	114	115	114
Sample Period	June 18- Sept. 9	June 20- Sept. 8	Feb. 29- Dec. 31	Mar. 4- Oct. 18	Apr. 4- Aug. 28	Apr. 4- Aug. 28	Apr. 4- Aug. 28

Table 5.4B
(continued)

[a]Rate-of-pay in 1974 for Irvine is a constant 32 cents per box since grove conditions are very nearly uniform.

[b]For orange pickers, there is no _a priori_ reason to think that grove attributes contributing to picking ease vary systematically by calendar date. Thus LDE is not a relevant explanatory variable and two-stage-least-squares estimating procedures were not necessary to avoid the collinearity problem between LDE and the environmental variables.

*Coefficient is significantly different from zero at the 0.10 level of the one-tailed t-test.

**Coefficient is significantly different from zero at the 0.05 level of the one-tailed t-test.

***Coefficient is significantly different from zero at the 0.025 level of the one-tailed t-test.

(The numbers in parentheses refer to the standard errors of the estimated coefficients).

Table 5.5A

Earnings Estimates by Two-Stage-Least-Squares for Lemon Pickers When LTM is Deleted. Dependent Variable = LI_t.

Picker: Variable	Upland 1 (1974)	Upland 1[a] (1973)	Upland 2 (1974)	Upland 2[a] (1973)	Upland 3 (1973)	Upland 4 (1973)	Upland 22 (1974)	Santa Paula 10[a] (1973)	Santa Paula 10 (1974)	Santa Paula 11 (1974)	Santa Paula 11[a] (1973)
Constant	1.410	1.177	2.731	2.620	0.662	-2.226	0.029	0.862	0.736	0.046	0.029
	(1.020)	(1.515)	(1.109)	(1.610)	(1.424)	(1.479)	(1.394)	(2.741)	(0.446)	(0.217)	(1.090)
Lw	-0.772	-0.159	-0.640	-0.448	-0.517	1.606	0.816	-0.291	-0.112	-0.428	-0.850
	(0.411)	(0.458)	(0.404)	(0.472)	(0.469)	(0.409)	(0.586)	(0.610)	(0.162)	(0.160)	(0.881)
LH	1.080	1.088	1.035	1.020	1.247	0.963	1.074	0.987	1.070	1.124	1.140
	(0.039)	(0.041)	(0.044)	(0.045)	(0.041)	(0.047)	(0.049)	(0.103)	(0.023)	(0.020)	(0.057)
LFR	0.260	0.041	-0.035	-0.127	0.188	0.107	-0.189	0.426	0.053	0.130	0.095
	(0.214)	(0.291)	(0.221)	(0.335)	(0.329)	(0.329)	(0.342)	(0.352)	(0.129)	(0.110)	(0.295)
LBT	-0.201	-0.002	-0.109	0.124	-0.048	0.342	0.239	-0.325	-0.114	-0.158	-0.200
	(0.111)	(0.103)	(0.111)	(0.005)	(0.105)	(0.012)	(0.160)	(0.102)	(0.040)	(0.028)	(0.149)
LOZ								-0.256***	-0.038***	-0.088***	-0.072
								(0.110)	(0.015)	(0.034)	(0.069)
LOZH	-0.021	-0.026**	-0.058***	-0.046**	-0.032**	-0.031**	-0.050**				
—2	(0.020)	(0.018)	(0.021)	(0.019)	(0.017)	(0.013)	(0.026)			0.894	
R	0.856	0.826	0.814	0.824	0.857	0.888	0.822	0.619	0.887	0.163	0.916
S.E.	0.208	0.244	0.210	0.286	0.211	0.301	0.256	0.304	0.169	[illegible]	0.186
D-W	2.090	1.89	1.81	1.470	1.901	1.704	1.827	1.67	1.930	1.42	1.470
F	185.190	159.69	136.84	171.860	162.63	137.842	126.644	14.14	327.020	484.79	104.390
Sampl Size	162	168	156	189	136	101	143	58	293	264	54
SPeriod	Apr. 1- Nov. 2	Mar. 17- Dec. 21	Apr. 1- Nov. 2	Mar. 19- Dec. 21	Mar. 17- Dec. 20	Mar. 16- Dec. 21	Apr. 17- Nov. 2	May 14- July 11	Jan. 3- Nov. 20	Mar. 3- Dec. 4	May 13- July 11

Table 5.5A
(continued)

[a] The results reported for this picker have been derived from various empirical testing procedures and regression specifications.

*Coefficient is significantly different from zero at the 0.10 level of the one-tailed t-test.

**Coefficient is significantly different from zero at the 0.005 level of the one-tailed t-test.

***Coefficient is significantly different from zero at the 0.025 level of the one-tailed t-test.

(The numbers in parentheses are the standard errors of the estimated coefficients).

Table 5.5B

Earnings Estimates by Ordinary-Least-Squares for Orange Pickers When LTM is Deleted.[b] Dependent Variable = LI_t.

Picker: Variable	Upland 2 (1973)	Upland 4 (1973)	San Bern. 5 (1973)	San Bern. 7 (1973)	Irvine 38 (1974)	Irvine 39 (1974)	Irvine 40 (1974)
Constant	-0.401	-0.885	1.018	-1.859	1.092	0.861	0.777
	(0.520)	(0.331)	(0.441)	(0.868)	(0.131)	(0.139)	(0.139)
Lw	0.420	0.442	0.119	1.230	a	a	a
	(0.272)	(0.171)	(0.178)	(0.260)			
LH	1.386	1.132	0.884	1.022	1.027	1.148	1.186
	(0.110)	(0.069)	(0.076)	(0.094)	(0.067)	(0.061)	(0.065).
LTR			-0.164	0.082			
			(0.110)	(0.165)			
LBT			-0.043	0.150			
			(0.051)	(0.095)			
LØZ	0.001	-0.062**			-0.108**	-0.031	0.049
	(0.057)	(0.036)			(0.059)	(0.061)	(0.057)
LØZH			-0.059***	-0.071**			
			(0.028)	(0.045)			
$\bar{R}^2$	0.752	0.856	0.487	0.800	0.680	0.759	0.752
S.E.	0.183	0.112	0.409	0.411	0.313	0.331	0.304
D-W	1.160	1.942	1.794	1.497	1.283	1.553	1.091
F	57.540	99.016	29.855	146.035	117.723	176.164	167.960
Sample Size	57	54	163	152	114	115	114
Sample Period	June 18- Sept. 9	June 20- Sept.. 8	Feb. 29- Dec. 31	& r. 4- Oct. 18	Apr. 4- Aug. 28	Apr. 4- Aug. 28	Apr. 4- Aug. 28

Table 5.5B
(continued)

[a]Rate-of-pay in 1974 for Irvine is a constant 32 cents per box since grove conditions are very nearly uniform.

[b]For orange pickers, there is no a priori reason to think that grove attributes contributing to picking ease vary systematically by calendar date. Thus LDE is not a relevant explanatory variable and two-stage-least squares estimating procedures were not necessary to avoid the collinearity problem between LDE and the environmental variables.

*Coefficient is significantly different from zero at the 0.10 level of the one-tailed t-test.

**Coefficient is significantly different from zero at the 0.05 level of the one-tailed t-test.

***Coefficient is significantly different from zero at the 0.025 level of the one-tailed t-test.

(The numbers in parentheses refer to the standard errors of the estimated coefficients).

coefficients but relatively significant temperature coefficients in Table 5.4. Of course, since we have no reason to suppose that temperature is an irrelevant explanatory variable and since temperature is obviously nonorthogonal to air pollution, only the air pollution'coefficients in Table 5.4 should be viewed as unbiased estimates. Nevertheless, since the air pollution coefficients of Table 5.5 generally do not exhibit major change from those in 5.4, one can tenatively and somewhat hesitantly conclude that, over the observed ranges of variation of the two variables, air pollution has greater relevance to picker earnings than does temperature.[3]

One must temper the generalizations of the above paragraph with the observation that the estimated equations for three orange pickers (Upland 2 (1973), Irvine 38 (1974), and Irvine 40 (1974)) have Durbin-Watson statistics probably indicative of negative autocorrelation of disturbances. It therefore seems likely that the estimated expressions for these individuals, as presented in Tables 5.4 and 5.5, are misspecified. More will be said shortly about the sources of this specification'error.

Finally, when reviewing Tables 5.4 and 5.5, the careful reader will have noted that the three grove attributes (BT, FR, TR) used by the packinghouses to determine the piece-work wage rate are frequently statistically non-significant and often seemingly of the wrong sign. However, because the piece-work wage rate is adjusted in accordance with changes in the values of these grove attribute variables, one cannot interpret their coefficients in the customary manner. Instead of representing the response of the picker's earnings to changes in the grove attributes variables, the coefficients represent the deviation in the individual picker's adjustment to the change from the adjustment to the grove attribute reflected in the piece-work wage rate. If this rate were always adjusted perfectly for the individual picker, the coefficients attached to the grove attribute variables would each be zero; that is, any change in one or more of the three grove attributes would have no effect whatsoever upon the picker's earnings.

Increases in Income Required to Compensate Pickers for Earnings Losses Due to Air Pollution. For expressions in which the temperature variable has been included, Table 5.6 below presents the calculated effects of air pollution upon the daily earnings of those lemon and orange pickers for whom

Table 5.6

Required Picker Income Compensations[a]

Picker: Statistic	Lemons Upland 2 (1973)	Lemons Upland 22 (1974)	Lemons Upland 3 (1973)	Lemons Santa Paula 10 (1973)	Lemons Santa Paula. 10 (1974)	Oranges San Bern. 5 (1973)	Oranges Irvine 38 (1974)
n	189	143	136	58	293	163	114
ΣI_t	2843.30	1645.93	1489.07	821.86	4063.85	1886.23	2063.40
$\Sigma \emptyset Z$				262.1			475.95
$\Sigma \emptyset ZH$	1251.30	1403.40	895.30			1064.06	
$\bar{I}_t$	15.34	11.51	10.95	14.17	13.87	11.57	18.1
$\overline{\emptyset Z}$				4.52	3.38		4.175
$\overline{\emptyset ZH}$	6.83	9.87	6.58			6.528	
$\hat{b}$	-0.038	,-0.047	-0.029	-0.242	-0.034	-0.061	-0.081
$\bar{V}$ in dollars	-0.090	-0.055	-0.048	-0.760	-0.140	-0.108	-0.351
$\bar{\bar{V}}$ in dollars	-0.280	-0.105	-0.144	-0.900	-0.380	-1.138	-0.446
$n\bar{\bar{V}}$ in dollars	'32.92	-15.02	-19.58	-52.20	-111.34	-185.49	-50.84
$\left(\frac{n\bar{\bar{V}}}{I_t+n\bar{\bar{V}}}\right)100$	-1.8%	-0.9%	-1.3%	-6.0%	-2.7%	-9.0%	- 2.5%

[a]These calculations were made from estimated expressions in which it was assumed the picker's work-time was institutionally fixed.

statistically significant air pollution coefficients have been obtained, A table showing the same calculations when the temperature variable has been excluded is not presented because the air pollution coefficient in these expressions is thought to be biased; that is, since temperature is known to be nonorthogonal to air pollution'and since we are unable to show that temperature is an irrelevant explanatory variable, calculations of required compensation using statistical results that do not account for temperature would be highly untrustworthy. The sole purpose of presenting estimates for expressions in which temperature is deleted has been to provide the reader a sense of the extent to which collinearity between temperature and air pollution affects the standard error and thus the statistical significance of the air pollution coefficient in estimated expressions including both variables.

Table 5.6 does not include pickers with statistically insignificant air pollution coefficients because we are able to reject, for these pickers only, and only within the context of the particular empirical specification, the hypothesis that air pollution influenced their earnings. For those pickers exhibiting statistically significant air pollution coefficients, the calculated losses represent, in accordance with the analytical construct presented in Chapter 4, the compensation the picker requires to make him indifferent between the presence or absence (except for "background" levels) of photochemical oxidants, given that he works as long as his picking crew. In Table 5.6, $n\bar{\bar{V}}$ represents this total required compensation for the picker during the period of observation and the bottom row of figures shows this required compensation as a percentage of what the picker's earnings would have been in the absence of air pollution. Thus, for example, in the 293 lemon groves in which Santa Paula 10 picked from January 3, 1974, to November 20, 1974, he required in compensation 38 cents per grove that he picked, $111.34 in total, and 2.7% of what his income would have been in the absence of air pollution.

Table 5.6 actually contains two calculations of picker's required compensations. Both assume that the response of the picker's earnings to variations In air pollution is a constant. The first calculation, 3, is

$$\bar{V} = \hat{b}\,\frac{\text{arithmetic mean of } I_t}{\text{arithmetic mean of } \emptyset Z \text{ or } \emptyset ZH} \equiv \hat{b}\,\frac{\bar{I}_t}{\overline{\emptyset Z} \text{ or } \overline{\emptyset ZH}},$$

where $\hat{b}$ is the estimated coefficient of the air pollution variable. This calculation gives the picker's required income compensation for the average grove he picks. The second calculation, $\bar{\bar{V}}$, is the picker's required income compensation per grove that he picked during the period of observation. It is:

$$\bar{\bar{V}} = \frac{\hat{b}}{n} \sum_{i=1}^{n} \frac{I_{t_i}}{\text{ith air pollution observation}},$$

where the i subscript indexes the groves in which the worker picked and n is the number of groves. Only the dollar magnitudes and percentages associated with $\bar{\bar{V}}$ are presented in Table 5.6 because $\bar{\bar{V}}$ takes greater account of the peculiarities of each grove in which the picker has worked. Calculations of required compensations that use $\bar{V}$ rather than $\bar{\bar{V}}$ will obviously give lower dollar and percentage magnitudes.

Does Air Pollution Impact Vary with the Picker's Physical Condition? The picking of citrus fruit is a physically strenuous activity, giving reason to speculate that over relatively long work-days the picker will become fatigued and therefore be more susceptible to the deleterious effects of air pollution.[4] However, the results reported in Tables 5.4 and 5.5 reflect the impact of air pollution on the earnings of various pickers for a wide range of work-day lengths. By including this entire range of work-day lengths in the sample used for each estimate in Tables 5.4 and 5.5, one obtains air pollution coefficients representing weighted averages of the picker's responses over all work-day lengths. In the absence of further analysis, it is impossible to disentangle the separate contribution to these weighted averages of assorted work-day lengths. Furthermore, it could be that the failure of the procedures used in Tables 5.4 and 5.5 to consider the differential effect of hours worked upon air pollution impact may, for certain pickers, have incorrectly resulted in statistical rejection of the hypothesis that air pollution influences picker earnings.

The results exhibited in Table 5.7 are the air pollution coefficients obtained by running the exact specifications of Tables 5.4 and 5.5 on partitionings by hours worked of the'identical observations of picker performance used in Tables 5.4 and 5.5. No Irvine workers are included in Table 5.7. They will be discussed separately. The results included in Table 5.7A include temperature as an explanatory variable, while those in Table 5.78 do not. Although the covariance F-test for single coefficients developed by Tiao and Goldberger (1962) could be used to test for statistically significant differences in the air pollution coefficients across partitionings, time did not permit the completion of this task for this report. A glance at the differences in magnitude among many of the coefficients for single pickers nevertheless leaves little doubt that statistically significant differences are fairly common.

The speculation that picker responsiveness to air pollution increases directly with the length of the work-day receives some support from the results exhibited in Table 5.7A.[5] For ten of the fifteen pickers in the table, the air pollution coefficient, an elasticity coefficient, increases in negative magnitude with increases in the length of work-day. Indeed, given the near-universal lack of significance (Upland 4 (1973) in lemons is the sole exception) of the air pollution coefficients in the partitionings representing relatively short work-days., it is tempting to assert that air pollution has little, if any, impact unless the work-day is in excess of about six or seven hours. The results of Table 5.7A are at least consistent with this interpretation.

The apparent tendency of air pollution impact to increase with increased work-day length is associated with the most interesting and important feature of Table 5.7A: with the exceptions of Upland 2 (1973) in oranges and Upland 4 (1973) in oranges, all pickers for whom statistically non-significant air pollution coefficients were obtained in Table 5.4 now have significant air pollution coefficients for the work-day partitioning greater than or equal to seven hours. In fact, it is plausible that the failure of Upland 4 (1973) in oranges to be significant is due to collinearity between temperature and air pollution. Note that in Table 5.7B, where temperature

Table 5.7A

Air Pollution Coefficients (and Standard Errors) for H Partitionings. Dependent Variable = LI_t.

Picker	Fruit	Statistic	$2.0<H<4.0$	$4.0<H<7.0$	$2.0<H<7.0$	$H \geq 7.0$
Upland 1 (1973)	Lemons	$\hat{b}_{LØZH}$	-0.001	-0.026		-0.096***
		s	(0.030)	(0.042)		(0.043)
		n	44	64		60
Upland 1 (1974)	Lemons	$\hat{b}_{LØZH}$	-0.008	0.019		-0.081***
		s	(0.050)	(0.039)		(0.036)
		n	37	63		62
Upland 2 (1973)	Lemons	$\hat{b}_{LØZH}$	-0.031	-0.065		-0.075
		s	(0.041)	(0.052)		(0.063)
		n	67	45		67
Upland 2 (1974)	Lemons	$\hat{b}_{LØZH}$	-0.043	0.080		-0.085**
		s	(0.053)	(0.044)		(0.044)
		n	31	63		62
Upland 2 (1973)	Oranges	$\hat{b}_{LØZ}$			0.181	0.030
		s			(0.188)	(0.087)
		n			14	43
Upland 3 (1973)	Lemons	$\hat{b}_{LØZH}$	-0.024	-0.048**		0.065
		s	(0.032)	(0.024)		(0.070)
		n	37	47		52
Upland 4 (1973)	Lemons	$\hat{b}_{LØZH}$	-0.035**	0.025		-0.137***
		s	0.022	(0.043)		(0.053)
		n	33	37		31
Upland 4 (1973)	Oranges	$\hat{b}_{LØZ}$			0.039	-0.005
		s			(0.066)	(0.065)
		n			14	40
Upland 22 (1974)	Lemons	$\hat{b}_{LØZH}$			-0.018	-0.061**
		s			(0.051)	(0.033)
		n			49	48

Table 5.7A
(continued)

Picker	Fruit	Statistic	2.0<H<4.0	4.0<H<7.0	2.0≤H<7.0	H>7.0
Santa Paula 10 (1973	Lemons	$\hat{b}_{LØZ}$ s n			-0.021 (0.288) 30	-0.299* (0.201) 28
Santa Paula 10 (1974	Lemons	$\hat{b}_{LØZ}$ s n	-0.047 0.045 60	-0.024 0.021 117		-0.020 (0.021) 116
Santa Paula 11 (1973	Lemons	$\hat{b}_{LØZ}$ s n			-0.271 (1.066) 30	-0.108** (0.053) 24
Santa Paula 11 (1974	Lemons	$\hat{b}_{LØZ}$ s n	0.075 (0.082) 34	-0.020 (0.023) 97		-0.055** (0.030) 90
San Bern. 5 (1973)	Orange	$\hat{b}_{LØZH}$ s n	-0.045 (0.049) 38	-0.118** (0.068) 46		-0.027 (0.081) 79
San Bern. 7 (1973)	Oranges	$\hat{b}_{LØZH}$ s n			0.032 0.069 57	-0.062*** 0.030 95

*Coefficient is significantly different from zero at the 0.10 level of the one-tailed t-test.

**Coefficient is significantly different from zero at the 0.05 level of the one-tailed t-test.

***Coefficient is significantly different from zero at the 0.025 level of the one-tailed t-test.

Table 5.7B

Air Pollution Coefficients (and Standard Errors) for 11 Partitionings When LTM is Deleted. Dependent Variable = LI_t.

Picker	Fruit	Statistic	$2.0 \leq H<4.0$	$4.0<H<7.0$	$2.0<H<7.0$	$H \geq 7.0$
Upland 1 (1973)	Lemons	$\hat{b}_{LØZH}$ s n	0.015 0.037 44	-0.047 (0.037) 64		-0.097*** (0.044) 60
Upland 1 (1974)	Lemons	$\hat{b}_{LØZH}$ s n	0.006 (0.047) 37	-0.025 (0.029) 63		-0.054** (0.030) 62
Upland 2 (1973)	Lemons	$\hat{b}_{LØZH}$ s n	-0.046* (0.035) 67	-0.099** (0.052) 45		-0.052 (0.045) 67
Upland 2 (1974)	Lemons	$\hat{b}_{LØZH}$ s n	-0.038 (0.046) 31	-0.038 (0.037) 63		-0.103*** (0.045) 62
Upland 2 (1973)	Oranges	$\hat{b}_{LØZ}$ s n			0.011 (0.140) 14	0.037 (0.068) 43
Upland 3 (1973)	Lemon5	$\hat{b}_{LØZH}$ s n	-0.038 -0.033 37	-0.034* (0.025) 47		0.038 (0.041) 52
Upland 4 (1973)	Lemons	$\hat{b}_{LØZH}$ s n	-0.047*** (0.020) 33	-0.001 (0.044) 37		-0.055* (0.036) 31
Upland 4 (1973)	Oranges	$\hat{b}_{LØZ}$ s n			0.009 (0.060) 14	-0.067* (0.044) 40
Upland 22 (1974:	Lemons	$\hat{b}_{LØZH}$ s n			-0.010 (0.046) 49	-0.031 (0.056) 48

Table 5.7B
(continued)

Picker	Fruit	Statistic	$2.0 \leq H < 4.0$	$4.0 < H < 7.0$	$2.0 \leq H < 7.0$	$H > 7.0$
Santa Paula 10 (1973)	Lemons	$\hat{b}_{L\O Z}$ s n			-0.074 (0.224) 30	-0.275* (0.180) 28
Santa Paula 10 (1974)	Lemons	$\hat{b}_{L\O Z}$ s n	-0.052 (0.045) 60	-0.024 (0.021) 117		-0.022 (0.021) 116
Santa Paula 11 (1973)	Lemons	$\hat{b}_{L\O Z}$ s n	0.020 0.070 34	-0.016 0.022 97		-0.055** (0.030) 90
Santa Paula 11 (1974)	Lemons	$\hat{b}_{L\O Z}$ s n			-0.045 (0.125) 30	-0.108*** (0.054) 24
San Bern. 5 (1973)	Oranges	$\hat{b}_{L\O ZH}$ s n	-0.035 (0.040) 38	-0.131*** 0.061 46		-0.038 (0.061) 79
San Bern. 7 (1973)	Oranges	$\hat{b}_{L\O ZH}$ s n			0.036 (0.065) 57	-0.069*** (0.030) 95

*Coefficient is significantly different from zero at the 0.10 level of the one-tailed t-test.

**Coefficient is significantly different from zero at the 0.05 level of the one-tailed t-test.

***Coefficient is significantly different from zero at the 0.025 level of the one-tailed t-test.

has been deleted from the estimated expression, the air pollution coefficient is fairly significant for the longer work-day partitioning.

In Table 5.8 are presented the estimates obtained for work-day partitionings of the three Irvine orange pickers. The partitioning of work-day lengths for these pickers causes the air Pollution coefficients for the shorter work-days to be statistically significant. This is diametrically opposed to the observed tendency for most other pickers of air pollution impacts to increase with increasing work-day lengths. In fact, the overall statistical results for the shorter work-day length estimates accord rather closely to those obtained for other pickers. The magnitudes and signs of the coefficients for each variable are similar to those obtained for other pickers, the Durbin-Watson statistic is very close to 2.00, and the F-values for the entire expression are highly significant. The estimates for these shorter work-days thus seem quite reliable. For the longer work-days, reliability must be sought elsewhere. The Durbin-Watson statistics imply negative autocorrelation and F-values for the entire expression are statistically nonsignificant. Plots of the residuals against the values of the dependent variable displayed a classic case of heteroscedasticity, Clearly, the statistical estimation procedure employed for this longer work-day length partitioning must be found wanting. With some consternation, we violated our data-grubbing ethic, and, now including a calendar date variable, we re-estimated the same expression for the longer work-days of these three Irvine pickers. The coefficient for this variable was nonsignificant; it added extremely little to the total explanation of the variations in the dependent variable; and neither the coefficients nor the standard errors of other variables were altered in any more than a minor way. Upon plotting the residuals of the estimates against time, *however*, a sine-curve pattern could be distinctly discerned. The period between each peak of the wave was consistently about two weeks long. We have no explanation for this phenomenon, nor do we understand why the apparent statistical quality of the estimates for the shorter and the longer work days should be so utterly different.

Table 5.8

Earnings Estimates by Ordinary: Least Squares for H Partitionings of Irvine Orange Pickers. Dependent Variable = LI_t.

Picker:	Irvine 38 (1974)		Irvine, 39 (1974)		Irvine 40 (1974)	
Variable	$2.0 \le H < 7.0$	$H \ge 7.0$	$2.0 \le H < 7.0$	$H \ge 7.0$	$2.0 \le H < 7.0$	$H \ge 7.0$
Constant	3.086 (1.512)	2.104 (1.270)	2.223 (1.767)	1.276 (1.503)	1.959 (1.288)	1.700 (1.524)
Lw	a	a	a	a	a	a
LH	1.086 (0.130)	0.316 (0.257)	1.210 (0.115)	0.581 (0.263)	1.176 (0.119)	0.898 (0.272)
LØZ	-0.143** (0.077)	0.080 (0.063)	-0.100* (0.068)	0.027 (0.065)	-0.171** (0.094)	0.164 (0.065)
LTM	-0.565 (0.395)	-0.038 (0.325)	-0.390 (0.472)	0.109 0.367	-0.266 (0.339)	-0.176 (0.370)
$\bar{R}^2$	0.612	0.014	0.705	0.031	0.701	0.176
S.E.	0.325	0.249	0.382	0.280	0.272	0.293
D-W	1.95	1.04	1.92	1.27	1.91	1.43
F	40.563	1.44	83.915	2.480	55.714	1.76
Sample Size	53	61	51	63	47	66
Sample Period	Apr. 4- Aug. 28	Apr. 4- Aug. 28	Apr. 4- Aug. 28	Apr. 4- Aug. 28	Apr. 4- Aug. 28	Apr. 4- Aug. 28

[a] Rate-of-pay is a constant 32 cents per box since grove conditions are very nearly uniform.

*Coefficient is significantly different from zero at the 0.10 level of the one-tailed t-test.

**Coefficient is significantly different from zero at the 0.05 level of the one-tailed' t-test.

***Coefficient is significantly different: from zero at the 0.025 level of the one-tailed t-test.

(The numbers in parentheses refer to the standard errors of the estimated coefficients).

In spite of our puzzlement with respect to the longer work-day estimates for the Irvine orange pickers, the fact still remains that a partitioning by work-day lengths for these pickers did result in statistically significant air pollution coefficients for the shorter work-day partitioning for each picker. If the reader is willing to take one instance of a statistically significant air pollution coefficient for each picker in either Table 5.4A, 5.5A, 5.7A, and/or 5.8 as being acceptable evidence of a deleterious air pollution impact upon a picker, then the proposition that sixteen of eighteen, or eighty-nine percent, of the pickers studied appear to have been significantly and negatively impacted cannot be rejected. Only Upland 2 (1973) and Upland 4 (1973), both in oranges, refuse to yield negative and statistically significant air pollution coefficients. Remembering, however, that those whom we have called different pickers are often the same pickers picking a different crop or in a different year, twelve of twelve, or one hundred percent, of the pickers studied appear to have had their work performance damaged by air pollution for at least one of the two crops in both the years studied. It would place some strain upon one's credulity to insist that this observed frequency of deleterious air pollution impacts across individuals is simply due to chance, particularly when it is recognized that the collinearity between temperature and air pollution in Tables 5.4A, 5.5A, 5.7A, and 5.8 increases the standard error of the air pollution coefficient and thus reduces its statistical significance, without, of course, biasing the coefficient itself.

The calculations in Table 5.9 are performed in a manner identical to those in Table 5.6. Unless the longest work-day length partitioning for a picker did not yield a statistically significant air pollution coefficient, only the air pollution coefficients, the earnings observations, and the air pollution observations falling within the longest partitioning are used to calculate $\bar{V}$ and $\bar{\bar{V}}$. Otherwise the coefficients and observations for a lesser work-day length partitioning that did have a statistically significant air pollution coefficient arc used. ΣI_t, however, refers to all work-day lengths. The percentage in the last column of the table is tlirrefore defined in exactly the same manner as the last row in Table 5.6: it is the compensation, in terms of a percentage of what his total earnings would be in

Table 5.9

Required Picker Income Compensations Calculated Using
Results of H Partitionings

Picker	Partitioning	$\hat{b}_{LØZ}$ or $LØZH$	ΣI_t	$\bar{V}$	$\bar{\bar{V}}$	$n\bar{\bar{V}}$	$\left(\frac{n\bar{\bar{V}}}{I_t + n\bar{\bar{V}}}\right)100$
Upland 1 (1973)	H>7.0	-0.096	$3163.33	$-0.429	$-0.667	-$40.44	-1.3%
Upland 1 (1974)	H>7.0	-0.081	3468.22	-0.240	-3.036	-24.75	-0.7%
Upland 2 (1973)[a,b]	H>7.0	-0.075	2399.30	-0.287	-0.500	-33.77	-1.2%
Upland 2 (1974)	H>7.0	-0.085	2899.37	-0.206	-0.339	-21.13	-0.7%
Upland 3 (1973)[b]	4.0<H<7.0	-0.043	1489.07	-0.078	-.236	-10.86	-0.7%
Upland 4 (1973)[a]	H>7.0	-0.137	1213.50	-0.365	-0.719	-22.22	-1.8%
Upland 22 (1974)[b]	H>7.0	-0.061	1645.93	-0.038	-0.088	-4.51	-0.3%
Santa Paula 10 (1973)[b]	H>7.0	-0.299	821.86	-0.667	-1.086	-30.44	-3.5%
Santa Paula 10 (1974)[b]	H>7.0	-0.020	4063.85	-0.098	-0.159	-18.91	-0.5%
Santa Paula 11 (1973)	H>7.0	-0.108	1020.63	-0.647	-0.765	-18.59	-1.8%
Santa Paula 11 (1974)	H>7.0	-0.055	4331.51	-0.325	-0.406	-37.29	-0.9%
San Bern. 5 (1973)[b]	4.0<H<7.0	-0.118	1886.23	-0.049	-0.077	-6.14	-0.3%
San Bern. 7 (1973)	H>7.0	-0.062	3529.50	-0.097	-0.166	-16.20	-0.5%
Irvine 38 (1974)[b]	2.0<H<7.0	-0.143	2063.40	-0.529	-0.685	-36.31	-1.8%
Irvine 39 (1974)	2.0<H<7.0	-0.100	2313.10	-0.316	-0.418	-21.32	-0.9%
Irvine 40 (1974)	2.0<H<7.0	-0.171	2650.36	-0.697	-0.877	-41.22	-1.6%

[a] Lemons only.

[b] This picker had a statistically significant air pollution coefficient for the impartitioned estimates. See Table 5.6.

the absence of air pollution during the entire period of observation, the picker requires to make him indifferent to the presence of air pollution.

Upon taking the percentage required compensations for the pickers in Table 5.6, as well as the same compensations for those pickers in Table 5.9 who do not appear in Table 5.6, and then calculating an unweighted arithmetic mean over all eighteen pickers, one obtains a figure of 2.0 percent. Calculating this same mean for all twelve individuals yields a lesser required compensation of 1.3%, where each crop and/or year for each individual is weighted by n.

The partitionings in Tables 5.7A, 5.70, and 5.8 lack an analytical basis. They were selected to provide similar numbers of degrees of freedom across partitionings for most pickers. Moreover, as the careful. reader will have already noted from Table 5.1, they do not refer to actual work-day lengths but rather to work-day lengths in a particular grove. Thus a picker could conceivably have worked three hours in one grove and six hours in another on a given day, yet not have his actual work-day length appear as nine hours in our data. Instead, the three hours would be counted as an observation in the less than four hours partitioning', while the six hours would appear in the middle partitioning. This means, then, that the air pollution coefficients for' the lower and middle ranges in Table 5.7 are not, representative of the interaction between hours worked and air pollution impact since 11 does not represent actual work-day length. However, this problem is, trivial for the 'upper partitioning because, with only an extremely few exceptions involving no more than an hour, all 'observations in this upper, partitioning have a one-to-one correspondence between hours worked in a particular grove and the length of, the actual work-day. Of course, since it is likely the lower and middle partitionings for at least some pickers include hours toward the end of long actual work-days, the calculations of required picker compensations in Table- 5.7A are biased downward, i.e., actual required compensations are higher. Some idea of the magnitude of this downward bias is provided by comparing the calculated required compensations for pickers in Table 5.6 with the calculated required compensations for these same pickers in Table 5.9. For the pickers appearing in both tables, the percentage required Compensations in Table 5.6 exceed those appearing in Table 5.9 by factors of as little. as one-half (Upland 22 (1973)) and as great as thirty (San Bernhrdino (1973)).

Given the downward biases in the absolute magnitudes of the percentage required compensations of Table 5.9, it seems reasonable to conclude that the representative individual in this study would have had to receive two to three percent of what his income would be in the absence of air pollution to make him indifferent to the presence of the air pollution levels to which he was subjected. This statement applies only to circumstances in which the picker chose to work each day as many hours as he was institutionally allowed.

Footnotes: Chapter 5

1. These differences among crop years could have been accounted for by a dummy variable. However, this would have presumed that the difference was explained solely by a shift in the intercept of the earnings expression. The responsiveness of earnings to air pollution would have implicitly been assumed to have remained the same in the two years. Because of the large number of observations we had available on each picker's performance, the course we adopted appeared less restrictive.

2. In retrospect, a dummy variable approach might not have been the best way to handle this problem. The labor camp managers may be saying that Fridays and Mondays are separable and distinct blocks of time; that is, they cannot be embodied in the time constraint (7) of the picker's decision problem, but must be treated as additional time constraints. This would imply that separate earnings expressions should be estimated for each of these two days.

3. Although it could easily be a coincidence, it is worth noting that the negative magnitudes of the nir pollution coefficients for the three Irvine workers vary directly with the worker's ages. Irvine 38 is in his early forties, Irvine 39 is in his late twenties, and Irvine 40 is only eighteen years old.

4. There exists sound empirical evidence in the economics literature to support this notion of declining marginal productivity with respect to increases in the number of hours worked. Feldstein (1967), for example, in a study using British data finds that the elasticity of output with respect to hours substantially exceeds that with respect to men.

5. There exists an alternative explanation for the association between picker responsiveness to air pollution and long work-days: high air pollution levels and long work-days may themselves be associated. For example, long work-days may occur primarily in the summer months when high air pollution levels also occur. Although no attempt was made to calculate a simple correlation coefficient between crew work-day lengths and air pollution levels, a scanning of the data files for several workers made it appear that long work-days are distributed more-or-less rectangularly over the entire calendar year.

6. One interesting alternative strategy, which involves using first differences in the dependent variable, is presented in Ashenfelter and Heckman (1974).